



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA PODNIKATELSKÁ

FACULTY OF BUSINESS AND MANAGEMENT

ÚSTAV INFORMATIKY

INSTITUTE OF INFORMATICS

ZÁLOHOVÁNÍ DAT A DATOVÁ ÚLOŽIŠTĚ

DATA BACKUP AND DATA STORAGE

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

Milan Boháček

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. Jiří Kříž, Ph.D.

BRNO 2020

Zadání bakalářské práce

Ústav: Ústav informatiky
Student: **Milan Boháček**
Studijní program: Systémové inženýrství a informatika
Studijní obor: Manažerská informatika
Vedoucí práce: **Ing. Jiří Kříž, Ph.D.**
Akademický rok: 2019/20

Ředitel ústavu Vám v souladu se zákonem č. 111/1998 Sb., o vysokých školách ve znění pozdějších předpisů a se Studijním a zkušebním řádem VUT v Brně zadává bakalářskou práci s názvem:

Zálohování dat a datová úložiště

Charakteristika problematiky úkolu:

Úvod
Cíle práce, metody a postupy zpracování
Teoretická východiska práce
Analýza současného stavu
Vlastní návrhy řešení
Závěr
Seznam použité literatury
Přílohy

Cíle, kterých má být dosaženo:

Cílem práce je vytvoření návrhu řešení zálohování dat pro zefektivnění práce s uloženými daty a zajištění jejich vyšší bezpečnosti.

Základní literární prameny:

DOSEDĚL, T. Počítačová bezpečnost a ochrana dat. Brno: Computer Press, 2004. ISBN 80-25-0106-1.

NELSON, S. Pro data backup and recovery. New York: Distributed to the book trade worldwide by Springer Science+Business Media, 2011. ISBN 1430226625.

SODOMKA, P. Informační systémy v podnikové praxi. Brno: Computer Press, 2006. 352 s. ISBN 80-251-1200-4.

SOSINSKY, B. A. Mistrovství – počítačové sítě. Brno: Computer Press, 2010. 840 s. ISBN 978-8-251-3363-7.

STOPKA, M. Storage Area Network. Abclinuxu.cz [online]. 2010 [cit. 2013-03-05]. Dostupné z: <http://www.abclinuxu.cz/clanky/storage-area-network-1-uvod>.

Termín odevzdání bakalářské práce je stanoven časovým plánem akademického roku 2019/20

V Brně dne 29.2.2020

L. S.

doc. RNDr. Bedřich Půža, CSc.
ředitel

doc. Ing. et Ing. Stanislav Škapa, Ph.D.
děkan

Abstrakt

Tato bakalářská práce se zaměřuje na problematiku spojenou se zálohováním dat a datovými úložišti. Je rozdělena do několika částí, ve kterých se zabývá teoretickými východisky, analýzou současného stavu a návrhem vlastního řešení.

Abstract

This bachelor's thesis focuses on problematics associated with data backup and data storages. It is divided into several parts where it deals with theoretical basis, analyses of current state and the design of own solution.

Klíčová slova

cloud, datová úložiště, node, databázový server, zálohování dat

Key words

cloud, data storages, node, database server, data backup

Bibliografická citace

BOHÁČEK, Milan. *Zálohování dat a datová úložiště* [online]. Brno, 2020 [cit. 2020-04-28]. Dostupné z: <https://www.vutbr.cz/studenti/zav-prace/detail/127664>. Bakalářská práce. Vysoké učení technické v Brně, Fakulta podnikatelská, Ústav informatiky. Vedoucí práce Jiří Kříž.

Čestné prohlášení

Prohlašuji, že předložená bakalářská práce je původní a zpracoval jsem ji samostatně. Prohlašuji, že citace použitých pramenů je úplná, že jsem ve své práci neporušil autorská práva (ve smyslu Zákona č. 121/2000 Sb., o právu autorském a o právech souvisejících s právem autorským).

V Brně dne 17. května 2020

podpis studenta

Poděkování

Tímto bych chtěl poděkovat panu Ing. Jiřímu Kříži, Ph.D. za vedení mé bakalářské práce a za jeho cenné rady i čas. Dále bych chtěl poděkovat panu Ing. Petrovi Bělousovi za poskytnutí informací pro analýzu a v neposlední řadě patří mé díky rodině, přátelům a známým, kteří mi po celou dobu byli oporou.

OBSAH

OBSAH	5
ÚVOD	9
CÍLE PRÁCE, METODY A POSTUPY ZPRACOVÁNÍ.....	11
1 TEORETICKÁ VÝCHODISKA PRÁCE	12
1.1 Zálohování dat	12
1.2 Způsoby zálohování	13
1.2.1 Flash disk.....	13
1.2.2 Cloudové služby.....	13
1.2.3 Externí disk.....	14
1.2.4 Zálohovací software	15
1.2.5 NAS server	15
1.2.6 SAN.....	16
1.3 Druhy záloh.....	17
1.3.1 Plná záloha.....	18
1.3.2 Přírůstková záloha.....	18
1.3.3 Rozdílová záloha.....	19
1.4 Systém RAID.....	19
1.4.1 RAID 0	19
1.4.2 RAID 1	20

1.4.3	RAID 5	21
1.4.4	RAID 6	22
1.5	Databázové clustery	23
1.6	UPS.....	24
1.7	Ochrana dat před živelnými pohromami	24
1.8	Ochrana dat před zničením nebo krádeží	25
2	ANALÝZA SOUČASNÉHO STAVU	26
2.1	O společnosti.....	26
2.2	Poskytující služby MasterAPP.....	26
2.2.1	NÁVRH A DESIGN	26
2.2.2	Vývoj a testování	27
2.2.3	PŘÍSTUPNOST APLIKACÍ	27
2.2.4	Marketing a distribuce.....	27
2.3	Vybavení datového centra	27
2.3.1	Kapacita připojení	28
2.3.2	Přímý propoj s ostatními datovými centry	28
2.3.3	KVM over IP	29
2.3.4	Zabezpečení přístupu.....	29
2.3.5	Ekologické chlazení	29
2.3.6	Záloha elektrické energie	29
2.3.7	Ochrana před útoky	29

2.4	Poskytující služby MasterDC	30
2.4.1	Cloud / VPS	30
2.4.2	Fyzické servery	30
2.4.3	Software.....	30
2.4.4	Datový prostor	31
2.5	O aplikaci.....	31
2.6	Současný stav aplikace	31
2.7	Aktuální zálohování	32
2.8	Současné zapojení	32
3	VLASTNÍ NÁVRHY ŘEŠENÍ.....	33
3.1	Návrh nového zapojení.....	33
3.2	Přidání dalšího HAProxy proxy serveru.....	34
3.3	Přidání dalších controllerů.....	34
3.4	Přidání více instancí (nodů) databáze.....	34
3.4.1	Zálohy.....	35
3.4.2	Scale-out	35
3.4.3	Zvýšení výkonu.....	35
3.4.4	Převzetí master funkce při selhání	35
3.4.5	Zabezpečení	36
3.5	Software.....	36
3.5.1	BackupPC.....	36

3.5.2	Instalace	37
3.5.3	Obecná konfigurace	37
3.5.4	Přidání stroje a konkrétnější nastavení	39
3.6	Zálohování Databázového serveru	39
3.7	Zálohovací skript	39
3.8	Cloudové úložiště	42
3.9	Záloha pomocí snapshotů	42
3.10	Zhodnocení navrhovaného řešení	43
ZÁVĚR		44
SEZNAM POUŽITÝCH ZDROJŮ		45
SEZNAM POUŽITÝCH OBRÁZKŮ		47
SEZNAM POUŽITÝCH GRAFŮ		48
SEZNAM POUŽITÝCH TABULEK		49

ÚVOD

Data se postupem času stávají čím dál důležitějšími. V dnešní době jsou často tím nejcennějším, co jako jednotlivec či společnost vlastníme. Zabezpečení jejich dostupnosti v okamžiku potřeby by tedy mělo patřit mezi naše nejpřednější priority.

Jedním ze způsobů, jak pravděpodobnost dostupnosti dat pro případ potřeby zvýšit, je zálohování, což je proces vytváření záložních kopií původních dat. Krom samotného procesu zálohování je však potřeba brát v úvahu také funkčnost a zabezpečení jednotlivých záloh a zálohovacích médií (úložišť), tak i stav a zabezpečení originálních dat a médií (úložišť,) na nichž jsou uložena. Jedině tak může být určitá dostupnost tížených dat zaručena.

Cílem bakalářské práce je vytvořit návrh pro zálohování dat aplikace **Numbo Call Protect & Blocker** od firmy MasterAPP, která vznikla na jaře roku 2014, jako softwarová divize úspěšné datacentrové společnosti Master Internet.

Bakalářská práce bude systematicky rozdělena na část praktickou a teoretickou. V teoretické části využiji především benefitů analýzy. Na začátek teoretické části vysvětlím, v čem spočívá záloha a popíšu různé způsoby zálohování. Dále vysvětlím, jaké jsou druhy záloh a přesunu se na popis systémů RAID. V poslední části teoretických východisek práce krátce představím problematiku databázových klusterů, akumulátorových záložních zdrojů (UPS) a ochrany dat před živelnými pohromami nebo lidským faktorem.

Bude následovat analýza současného stavu, ve které jsem vycházel z informací poskytnutými zaměstnanci firmy, webovými stránkami firmy nebo z vlastního pozorování při osobních návštěvách firemních prostor. Osvětlím zde vybavení datového centra MasterDC Brno, které patří do vlastnictví firmy Master Internet a využívá ho mimo jiné k zálohování vlastních aplikací ze své softwarové divize MasterAPP. Stejně tak popíšu aktuální stav zálohování dat aplikace, pro kterou budu dělat návrh řešení a požadavky zákazníka na to, co by měl mnou vypracovaný návrh splňovat. V poslední části bakalářské práce, kterou je návrh vlastního řešení, se budu zabývat návrhem řešení

pro optimální a bezpečnou zálohu dané aplikace, a přitom se pokusím i vylepšit způsob, jakým aplikace bude přistupovat ke své databázi. Na konci této části také zmíním nějaké dodatečné způsoby zálohy, které by se daly k tomuto řešení použít.

CÍLE PRÁCE, METODY A POSTUPY ZPRACOVÁNÍ

Cílem práce je vytvoření návrhu řešení zálohování dat pro zefektivnění práce s uloženými daty a zajištění jejich vyšší bezpečnosti.

Proto, aby zmíněného cíle mohlo být dosaženo, budou jako první část práce vypracována teoretická východiska. Ty nám poslouží jako znalostní podklad pro další části, které se budou zabývat analýzou současného stavu a vlastním návrhem řešení.

Pro analýzu současného stavu budou použity empirické metody, a to rozhovory se zaměstnanci datacentra a vlastní pozorování. Vybavení datacentra a současný stav zálohování aplikace bude popsán na základě získaných informací a obdržených podkladů od zaměstnanců. Bude potřeba se seznámit se systémy a službami, které datacentrum nabízí a probrat se zaměstnanci různé možnosti, které by mohly být využity v rámci řešení zálohy.

1 TEORETICKÁ VÝCHODISKA PRÁCE

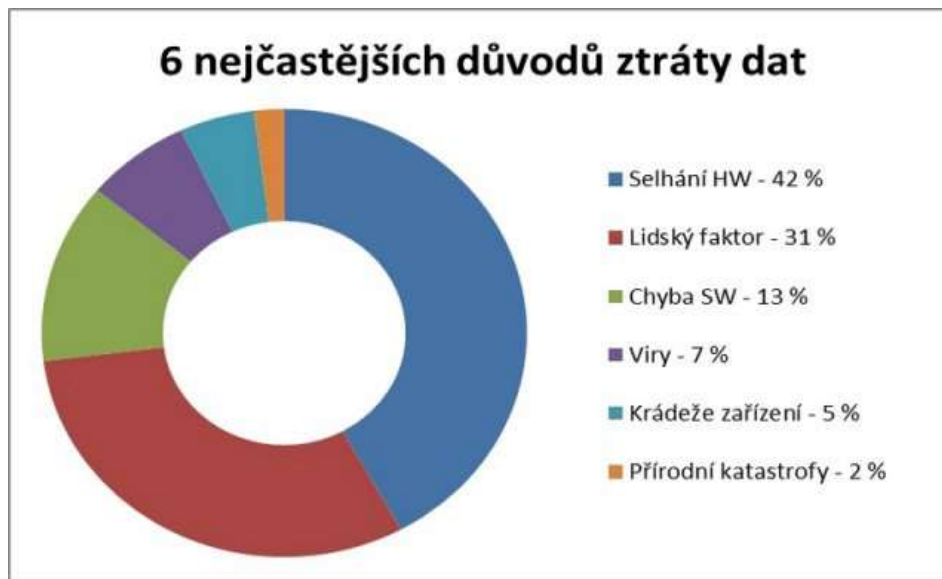
V této části se budeme zabývat teoretickými východisky. Ty nám poslouží jako znalostní podklad pro další části práce, které se budou zabývat analýzou současného stavu a vlastním návrhem řešení.

1.1 Zálohování dat

Snad každý v dnešní době používá nějaké elektronické zařízení, ať už jsou to počítače, digitální fotoaparáty, mobilní telefony nebo chytré hodinky. Všechna tato zařízení v sobě mají digitální informace (data), které jsou uloženy do obrovského počtu paměťových buněk. O tyto data lze velmi snadno přijít např. při poškození úložného zařízení nebo při lidském zásahu, a proto bychom data měly zálohovat (1).

Samotné zálohování je proces, kdy jsou vybraná data uložena na jiném médiu, než na kterém jsou aktuálně uložena. Vznikne tedy záložní kopie oněch dat. Poté v případě nedostupnosti původních dat, a to ať už z jakéhokoliv výše zmíněného důvodu, jsou data obnovena ze záložní kopie. Při jakékoliv obnově dat dochází k jejich ztrátě, minimálně těch, která byla uložena od poslední zálohy. Zálohování by se tedy mělo provádět pravidelně a co nejčastěji (2, s. 61).

Co se týče záložních kopií, tak ty by měly být uloženy odděleně od původních dat. Dané kopie pak měly být chráněny minimálně stejně dobře jako data původní a měla by se provádět pravidelná kontrola jejich funkčnosti pro případ potřeby obnovy (2, s. 62).



Obrázek 1: Důvody ztráty dat
(Zdroj: 11)

1.2 Způsoby zálohování

1.2.1 Flash disk

Jednou z nejpopulárnějších možností jak udržet data relativně v bezpečí, je zálohovat data na malý a vždy dostupný flash disk. Flash disk je po ruce téměř vždy a nahrát na něj data zvládne snad každý (3).

V dnešní době se kapacity USB flash disků pohybují v řádech desítek i stovek GB, takže doby, kdy na flashku nebylo možné zálohovat celé fotoalbum jsou nenávratně pryč. Dnes se na flash disk vejde několik desítek videí i ve 4K a tisíce dokumentů a prezentací (3).

1.2.2 Cloudové služby

Cloudové služby jako např. Google Drive, Dropbox nebo OneDrive jsou velmi dostupnou formou zálohování, kterou si oblíbili zejména lidé, kteří jsou často na cestách (3).

Velkou výhodou je právě to, že zálohovat lze z kavárny v centru města, z kanceláře nebo při cestě na dovolenou. Pro zálohu totiž stačí jakékoli připojení k internetu. Díky zrušení

vysokých roamingových poplatků můžete využívat své datové připojení i v zahraničí a jedná se o jedno z nejlepších řešení na malé soubory (3).

U cloudových záloh je důležité nastavení velmi bezpečného hesla. Heslo je to jediné, co drží vaše data v bezpečí a při zvolení špatného hesla jsou vaše data v podstatě veřejná a dostat se k nim není pro útočníka vůbec složité (3).

Dalším problémem může být omezená velikost úložiště. Využívání cloudu je totiž placené a za několik set korun je možné dosáhnout pouze stovek GB. Jestliže se nezaplatí pravidelná platba, poskytovatel cloudové služby vám data nenávratně smaže (3).

1.2.3 Externí disk

Druhou nejrozšířenější formou zálohy je externí disk. Rychlost externího disku je v dnešní době podobná již zmíněnému flash disku, avšak kapacita se běžně pohybuje v řádu terabajtů (3).

Externí pevné disky se snadno přenášejí, takže s sebou můžete nosit klidně obsah dat srovnatelného s obsahem v běžném počítači. Uživatelé také ocení snadné použití disku jako odkladiště dat, která na notebooku již nechťejí, jako například staré fotografie, filmy či dokumenty z minulých let (3).

V dnešní době jsou pevné disky nahrazovány technologií SSD. Jelikož SSD disky umějí pracovat s rychlostmi kopírování dat kolem 500MB/s což žádný flash disk neumí, jde určitě o nejrychlejší způsob externí zálohy (3).

Moderní SSD disky a pevné disky jsou navíc velice odolné a použitelné i na cestách, kde hrozí pád nebo jiný vliv prostředí jako vlhkost a prach (3).

Některé disky rovněž nabízejí automatický software, který zálohu nastaví a uživatel si tak nemusí pamatovat, že musí kopírovat data (3).

1.2.4 Zálohovací software

Zálohovací software umožňuje automaticky zálohovat data v počítači, ukládat různé verze dokumentů v různá časy a na různá místa (3).

Ideální je potom zálohování na více míst, což nové zálohovací software umějí. Každou noc se zálohují všechny dokumenty na externí disk, několikrát denně se udělá záloha na další počítače v síti a každý týden se udělá záloha do cloudu, například o víkendu (3).

Nejlepší řešení je zálohovat data na více míst zároveň, což novým zálohovacím softwarům nedělá žádný problém. Několikrát za den se dělá záloha na další počítače v síti, na konci dne se udělí záloha na externí disk a poté nejlépe na konci týdne se dělá záloha do cloudu (3).

Software lze snadno nastavit pro jakýkoli způsob zálohy, ať už se jedná o domácnosti nebo o firemní prostředí, cílové zařízení nerozhoduje. Některé zálohovací SW mají dokonce cloud v ceně, takže většinou stačí jen jednoduché nastavení ve většinou přehledném grafickém prostředí a je vystaráno (3).

1.2.5 NAS server

NAS server se ovládá prostřednictvím většinou jednoduchého webového rozhraní, protože k němu nelze připojit klávesnici, myš ani monitor. Jak můžeme odvodit z názvu, jedná se o síťové úložiště, tudíž je nejvhodnější způsob připojení do sítě pomocí síťového kabelu vedeného z routeru či switchu. NAS lze připojit také bezdrátově pomocí wifi, avšak zde může docházet ke snížení přenosové rychlosti dat (4).

Po úspěšné instalaci se vám NAS server hlásí jako zařízení v síti a přístup k němu je stejně jednoduchý jako na váš lokální disk (4).

NAS servery jsou navrženy tak, aby zvládaly nepřetržitý provoz a mohly tak plnit svou úlohu centrálního síťového úložiště, FTP serveru apod. I při svém neustálém provozu je jejich spotřeba elektrické energie velmi nízká. Jeho spotřeba se pohybuje mezi 10-40 W v závislosti na typu a v úsporném režimu, kdy disky nepracují tak ještě méně. NAS mají

také mnoho dalších možností jako například přehrávání filmů rovnou na TV, sdílení tiskárny nebo zapojení kamerového systému pro kancelář nebo byt. Dnes jde o jednu z nejoblíbenějších možností, jak sdílet a zálohovat data (4).

Zapojení NAS serveru do domácí sítě není vůbec nic složitého a za pomoci manuálu či průvodce zvládne instalaci v dnešní době téměř každý (4).



Obrázek 2: NAS
(Zdroj: 4)

1.2.6 SAN

SAN (Storage Area Network) je síť, která propojuje pracovní stanice a servery se zařízeními na ukládání dat. Taková síť je vyhrazena k přístupu k uloženým a zálohovaným datům a je zpravidla budována pomocí vysokorychlostních optických kabelů. Ve zkratce se tedy SAN používá k ukládání a ochraně dat (5).

SAN jsou běžně založeny na technologii Fibre Channel (FC), která využívá protokol Fibre Channel Protocol (FCP) pro otevřené systémy a proprietární varianty pro

mainframy. Použití technologie Fibre Channel oproti Ethernet (FCoE) navíc umožňuje přesunout provoz FC napříč existující vysokorychlostní ethernetovou infrastrukturou a konvergovat úložiště a IP protokoly na jeden kabel. Lze použít i jiné technologie, jako je Internet Small Computing System Interface (iSCSI), běžně používané v malých a středních organizacích jako levnější alternativa k FC, a InfiniBand, běžně používané ve vysoce výkonných výpočetních prostředích. Kromě toho je možné použít brány (gateways) k přesunu dat mezi různými technologiemi SAN (6).

Právě sdílené diskové úložiště přes síť je hlavní důvod proč používat SAN. SAN poskytuje rychlé přenosové médium pro přístup k datům, zálohování, obnovu a archivaci dat a umožňuje přesouvat data mezi různými systémy (5).

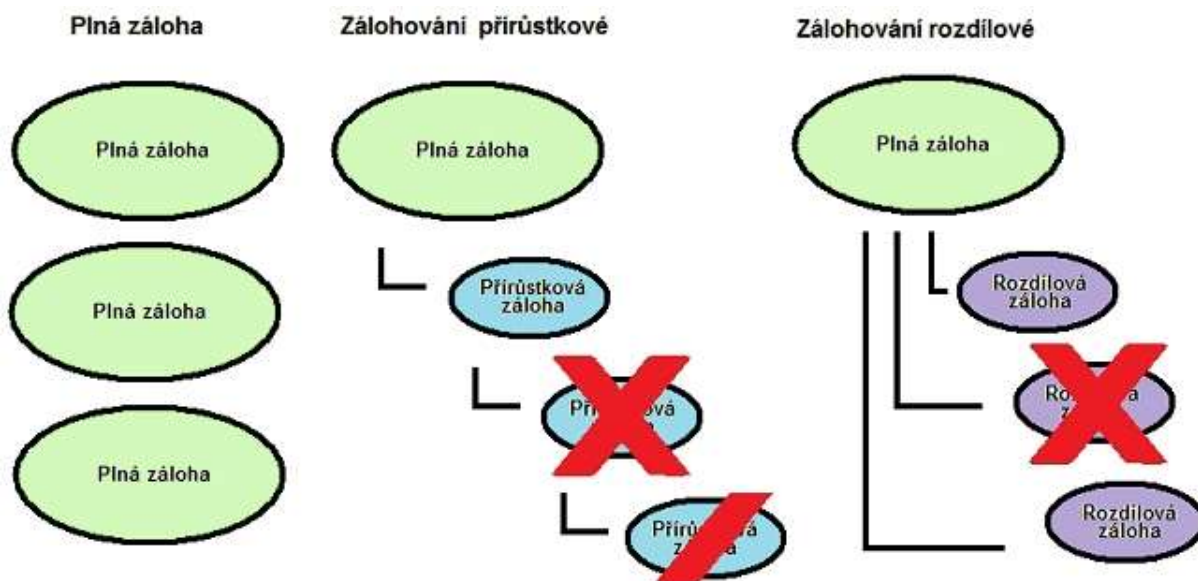
SAN se vyplatí ve velkých společnostech, které mají např. 25 a více serverů nebo společností, které potřebují velkou rychlost přístupu k datům. Síťové úložiště se také hodí v případě většího množství souborových serverů, zde je lepší NAS, nebo při potřebě zálohovat data delší dobu nebo např. streamovat video, zde se více hodí SAN (5).

SAN by se neměl implementovat v malé organizaci o pár serverech, která nemá problémy s výkonem databázových aplikací, zálohováním ani správou dat. Rozhodně se nevyplatí používat SAN pro aplikace jako webový server (pro ten je výhodný NAS), servery primární infrastruktury (DNS, WINS, doménové řadiče atd.), servery, které vystačí s méně než 100 GB dat (nemusí platit v případě, že je aplikace vysoce náročná na rychlost přístupu k datům) (5).

1.3 Druhy záloh

Používají se různé strategie zálohování pro různé podmínky. Zvolení si náležité strategie funguje v závislosti na tom, zda je třeba s určitými zálohami nutno pracovat častokrát či naopak, je nutností maximalizace délky archivace zálohovaných dat.

Používáme 3 typy zálohování: plná, přírůstková a rozdílová



Obrázek 3: Druhy záloh
(Zdroj: 8)

1.3.1 Plná záloha

Prostřednictvím metody plného zálohování je v mnoha případech zálohován celý počítač, a to včetně operačního systému. Vytváří se tak obraz disku. Podmínkou tohoto specifického typu zálohování je specializovaný software. Plné zálohování zabírá ze všech metod největší časový a prostorový úsek. Toto je důvodem pro fakt, že je nejčastěji využíván společně s metodou rozdílovou či přírůstkovou (7).

1.3.2 Přírůstková záloha

Záloha přírůstková, nebo jinak řečeno inkrementální, je typ zálohy, který zahrnuje pouze taková data, která byla od poslední plné přírůstkové zálohy nějakým způsobem pozměněna. Tím pádem nezabírá její tvorba tolik času jako metoda plného zálohování, avšak právě proto, že její součástí nejsou všechna data, je pro obnovení nutností mít k dispozici veškeré předcházející přírůstkové zálohy a také původní zálohu plnou (8).

1.3.3 Rozdílová záloha

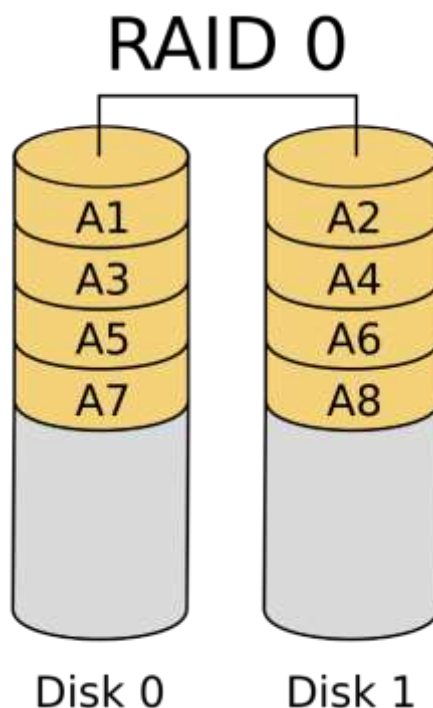
Metoda rozdílová či diferenciální je takový druh zálohy, jehož součástí jsou pouze ta data, která byla nějakým způsobem pozměněna od doby provedení plné zálohy. U předchozí příпустkové zálohovací metody je každý nově vytvořený záložní soubor součástí „řetězce.“ U metody rozdílového zálohování je naopak vytvořen zcela nezávislý soubor, jehož součástí jsou všechny změny provedené od vytvoření původní plné zálohy. Obecně je pravdou, že rozdílová metoda zálohování nezabírá tak velké množství času jako příпустková, jelikož při ní není nutností zpracování dlouhého řetězce předchozích záloh (9).

1.4 Systém RAID

RAID (Redundant Array of Independent Disks) – v překladu se jedná o vícenásobné diskové pole nezávislých disků. Je to metoda zabezpečení dat, která se provádí, aby se zamezilo selhání pevného disku. Pracuje na principu uložení dat na více nezávislých disků, a to tak, že uložená data jsou zachována i v případě poškození některého z výše zmíněných disků. Existuje možnost zvolení si z vícero úrovní zabezpečení. Metoda RAID je nejčastěji využívána na serverech, avšak nelze pomocí ní nahradit zálohování dat (7).

1.4.1 RAID 0

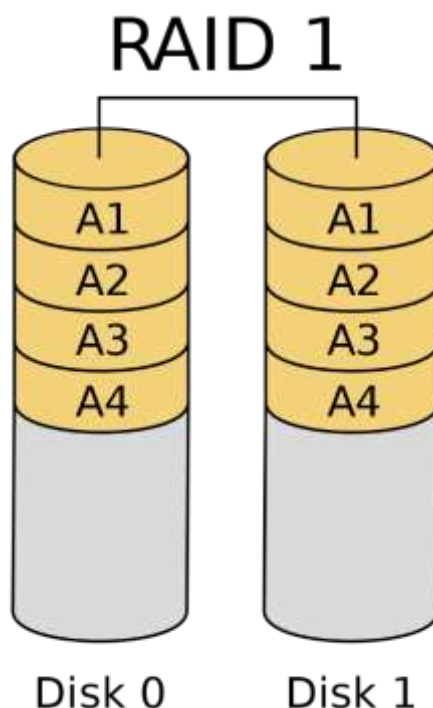
RAID 0 využívá techniku zvanou „striping,“ která rozděluje data na kousky a zapisuje je na větší množství disků, které jsou prezentovány jako jeden velký disk. Tato metoda zvyšuje pouze rychlost, která je měřena počtem zapojených disků. Nutností je mít zapojeny minimálně dva disky. Co se nevýhod v tomto zapojení týče, je důležité zmínit, že jakmile selže jeden z disků, všechna data jsou ztracena (11).



Obrázek 4: Raid 0
(Zdroj 12)

1.4.2 RAID 1

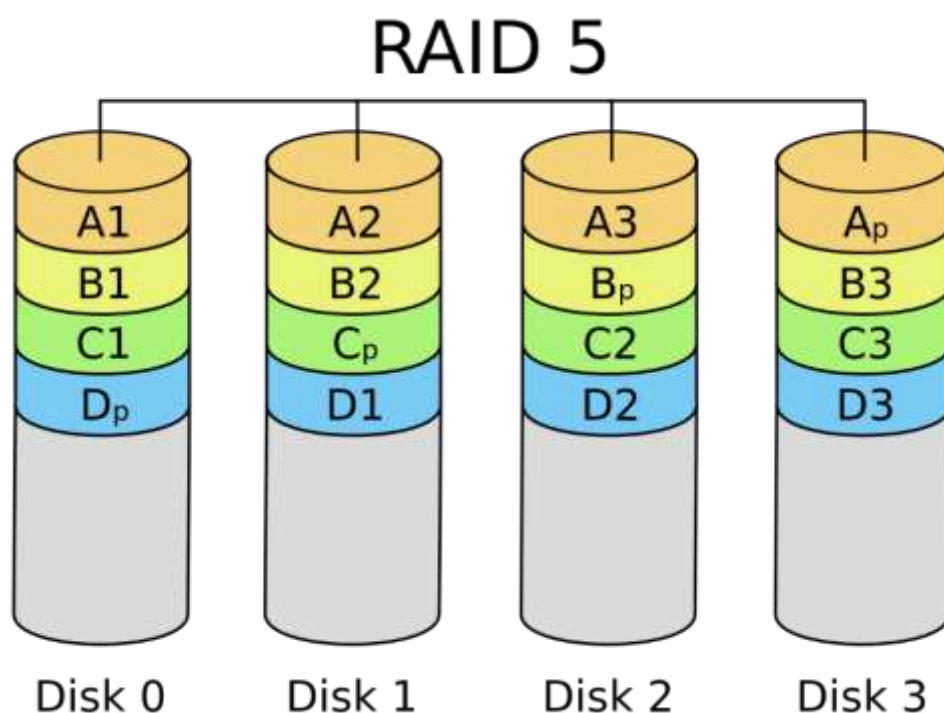
RAID 1 využívá techniku, která se nazývá „mirroring“, což je v překladu zrcadlení uložených dat. Jedná se o zálohu ve své nejběžnější formě, vyžaduje zapojení minimálně dvou disků. V podstatě jde o více disků, které obsahují ty samé informace a jakmile jeden z disků selže, zbylé ho dokáží obnovit. Jedinou nevýhodou jsou náklady. Je-li potřeba 1 TB úložného prostoru, je třeba obstarat dva 1 TB disky a jelikož jsou informace zapisovány na všechny disky současně, je rychlost zápisu rovná disku s nejmenší rychlostí (11).



Obrázek 5: Raid 1
(Zdroj: 12)

1.4.3 RAID 5

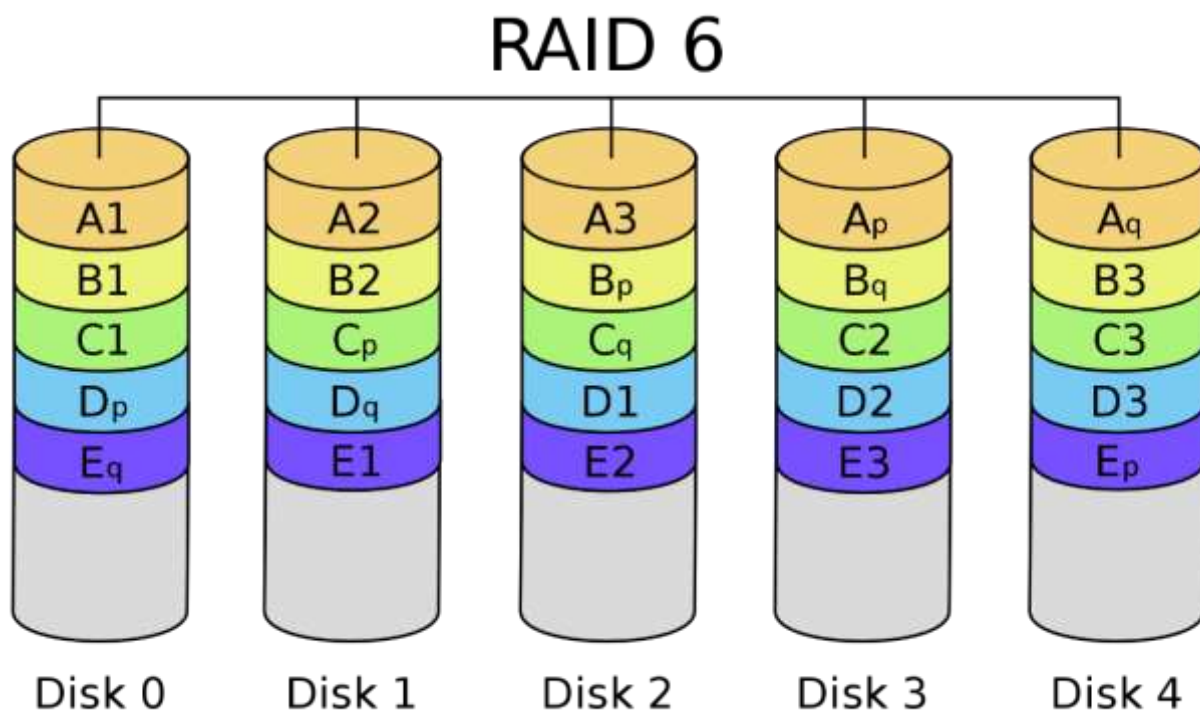
Tato úroveň RAID poskytuje lepší výkon než zrcadlení a odolnost proti chybám. U RAID 5 jsou data a parita (což jsou dodatečná data používaná pro obnovu) prokládána na třech nebo více discích. Pokud u disku nastane chyba nebo začne selhávat, data se z tohoto distribuovaného datového a paritního bloku znovu vytvoří, hladce a automaticky. Systém je v podstatě stále funkční, i při jednom nefunkčním disku a lze ho tak v klidu vyměnit. Další výhodou RAID 5 je to, že umožňuje mít mnoho NAS a serverových disků „vyměnitelných za provozu“, což znamená, že v případě selhání jednotky v poli, může být tato jednotka vyměněna za novou jednotku bez vypnutí serveru nebo NAS a bez nutnosti přerušit uživatele, kteří mohou přistupovat k serveru nebo NAS. Je to skvělé řešení pro odolnost proti chybám, protože až disky selžou, data mohou být přestavěna na nové disky (11).



Obrázek 6: Raid 5
(Zdroj: 12)

1.4.4 RAID 6

Tento typ je obdobou zapojení stejně jako u RAID 5, z toho důvodu, že na každém z disků využívá dva paritní bloky a paritní data jsou uložena střídavě na všech discích. Na rozdíl od RAID 5 je výhodou prevence vůči vypadnutí dvou disků, avšak stejnou rychlostí čtení a nižší rychlostí zápisu. Minimální počet potřebných disků na zapojení jsou čtyři (11).

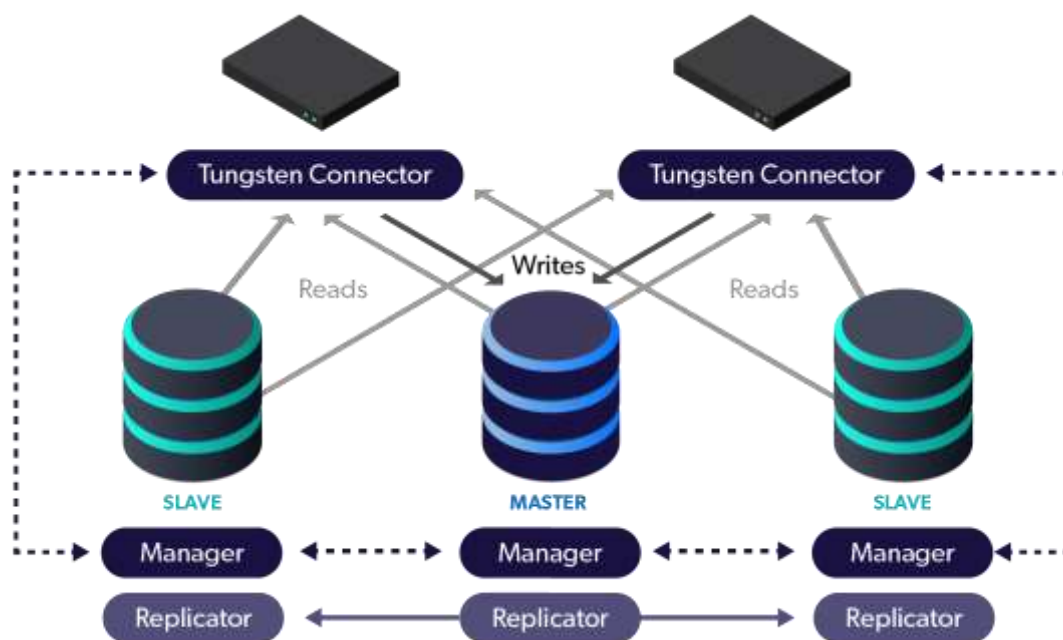


Obrázek 7: Raid 6
(Zdroj: 12)

1.5 Databázové clustery

Clustering databází je proces kombinování více než jednoho serveru nebo instancí připojení jedné databáze. Někdy jeden server nemusí být vhodný pro správu množství dat nebo počtu požadavků, to je situace, kdy je třeba datový klastr. Clustering databází, clustering serverů SQL a SQL clustering jsou úzce spojeny s SQL, což je jazyk používaný pro správu informací o databázi (13).

V clusterové architektuře jsou všechny požadavky rozděleny mezi více počítačů, takže individuální požadavek uživatele je zpracován a vytvořen několika počítačovými systémy. Clustering je použitelný díky schopnosti vyvažování zátěže a vysoké dostupnosti. Pokud se jeden uzel zhroutí, požadavek je zpracován jiným uzlem. V důsledku toho existuje jen málo nebo žádné možnosti absolutních selhání systému (13).



Obrázek 8: Databázový cluster
(Zdroj: 14)

1.6 UPS

UPS (Uninterruptible Power Supply), nebo v češtině akumulátorový záložní zdroj, je ve své podstatě zařízením, zapojeným mezi napájecí síť a zařízením, které má být napájeno. Jeho součástí je akumulátor, jenž je postupně dobíjen ze sítě. Vypadne-li síť, přebírá pak funkci napájení připojeného zařízení. Uživatel je na tuto skutečnost povětšinou upozorněn zvukovým signálem. Hlavním problémem je životnost a celková výdrž akumulátoru (2, s. 55 - 56).

1.7 Ochrana dat před živelnými pohromami

Bohužel, naše uložená data nejsou v nebezpečí pouze před nekalými živly. Jsou ohrožena také živelnými pohromami, které jsou v mnoha případech nepředvídatelné. Nejpodstatnějšími pohromami, co se poškození dat týče, jsou zemětřesení, povodeň či klimatické změny ve velké míře (2, s. 54).

Tudíž je nutné naše data ochraňovat i před těmito nežádoucími vlivy. Samozřejmostí je bezpečné umístění a ochrana prostorů, ve kterých se media s uloženými daty nacházejí. Další možností je samozřejmě pravidelné zálohování dat (2, s. 55).

1.8 Ochrana dat před zničením nebo krádeží

Co se uložení dat týče, nacházejí se na určitých datových nosičích, které mohou být buďto zcizeny či dokonce zničeny. Právě z tohoto důvodu je třeba dbát na zabezpečení prostor, ve kterých se tyto datové nosiče nacházejí (2, s. 52).

Způsobů, jak výše zmíněné prostory zabezpečit, existuje spousta. Příkladem může být kontrola oprávněnosti osob, kterým bude do prostor udělen přístup, a to nejlépe hned při vstupních dveřích do budovy. Dále existuje možnost vybavení oken detektory pohybu, monitorizace haly za pomoci bezpečnostních kamer a samozřejmě možnost uzamknout celé prostory, kde se nosiče nacházejí (2, s. 53).

2 ANALÝZA SOUČASNÉHO STAVU

V této kapitole se budu věnovat popisem firmy a aplikace, pro kterou řeším zálohování dat. Také zhodnotím současný stav zálohování v rámci této aplikace.

2.1 O společnosti

Vývojářské studio MasterAPP vzniklo na jaře 2014 jako softwarová divize úspěšné datacentrové společnosti Master Internet. Stabilní zázemí pro vývoj aplikací jim tak poskytují jejich vlastní datová centra MasterDC.

Master Internet se nachází v blízkosti centra města a jeho páteřních komunikací. MasterDC Brno provozuje všechny nabízené služby od housingu po cloud hosting. Plný provoz ve vlastních prostorách zahájili v roce 2005 a od té doby se stále rozšiřují.

2.2 Poskytující služby MasterAPP

Firma MasterAPP v Brně nabízí spoustu služeb od návrhu aplikací až po marketing, některé z nich tu tedy představím.

2.2.1 NÁVRH A DESIGN

- Analýza proveditelnosti
- Průzkum trhu
- Nacenění
- Koncept
- Návrh UX
- Wireframes
- Prototypy
- Design mobilní aplikace

2.2.2 Vývoj a testování

- Vývoj pro iOS
- Vývoj pro Android
- Vývoj pro Windows
- Backend
- Frontend
- Testování funkčnosti a bezpečnosti
- Uživatelské testování
- Audit přístupnosti

2.2.3 PŘÍSTUPNOST APLIKACÍ

- Pravidla přístupnosti WCAG 2.1
- Zákon č. 99/2019
- Audit přístupnosti
- Školení přístupnosti

2.2.4 Marketing a distribuce

- Překlad a lokalizace
- Publikace
- App Store Optimization (ASO)
- Search Engine Optimization (SEO)
- Marketingová propagace

2.3 Vybavení datového centra

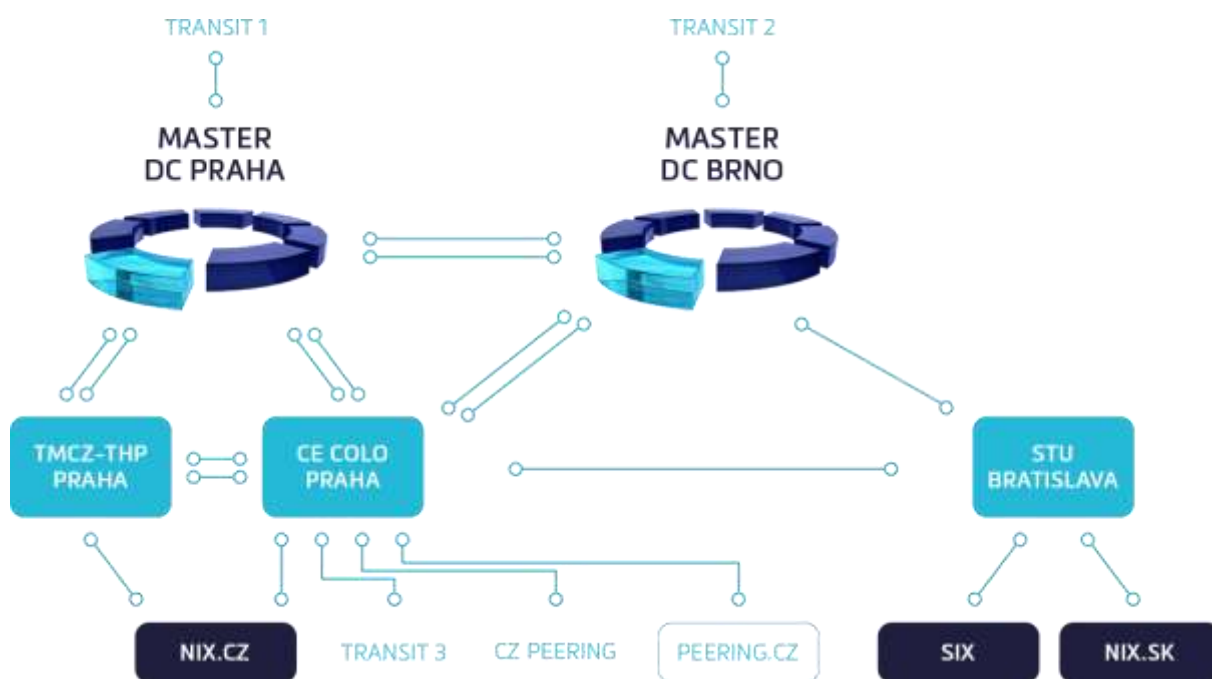
Vybavení datového centra Master DC Brno, ve kterém jsou data pro aplikaci zálohována je více než dostatečné. V této části poskytnu nějaké informace o vybavení, kterým datacentrum disponuje.

2.3.1 Kapacita připojení

Datové centrum je vybaveno několika 10 Gbit ethernetovými kabely, včetně dostatečné redundance pro zajištění bezproblémového a rychlého internetového připojení.

2.3.2 Přímý propoj s ostatními datovými centry

Datové centrum MasterDC Brno je přímo propojeno s ostatními datacentry, takže při výpadku jedné z linek se provoz automaticky přesměruje do záložního propoje, což zaručí vysokou dostupnost. Jsou použita prověřená zařízení a protokoly a spoléhá se pouze na kvalitní, moderní technologie (leaf-spine L3 design s použitím 100 GB Ethernetu). Konektivita je zprostředkována přes sítě Telia a RETN, a to s velkými českými operátory, jako O2 nebo UPC. Celá infrastruktura sítě je chráněna hardwary Radware DefensePro. Jsou to datacentra SIX, CeCoLo a jejich další pobočka v MasterDC Praha.



Obrázek 9: Propojení datacenter
(Zdroj: 15)

2.3.3 KVM over IP

Díky rozhraní KVM over IP se lze odkudkoliv připojit přes IP a získat tak vzdálený přístup ke svému zařízení. Tento způsob připojení poskytuje stejné možnosti jako při přímé práci s daným strojem. Vy tak můžete sedět třeba na druhém konci světa.

2.3.4 Zabezpečení přístupu

Před vstupem do datacentra musí jeden z členů support týmu přiložit dlaň k biometrické čtečce a nechat se identifikovat podle krevního řečiště. Bezpečí serverů je prioritou a vedle nejlepších technologií ho 24/7 zajišťuje support tým i fyzická ostraha.

2.3.5 Ekologické chlazení

MasterDC chladí free coolingem a klimatizačními jednotkami s nižší spotřebou energie. Na životní prostředí myslí i v krizových situacích, a proto funguje jejich protipožární systém na bázi plynového hašení s šetrným ekologickým plynem FM 200.

2.3.6 Záloha elektrické energie

Transformátory s několika zálohovanými zdroji elektřiny jsou schopny poskytovat elektrickou energii během kratších výpadků. Pro překlenutí těch delších, datacentrum disponuje dieslovými motorgenerátory, které bez tankování vydrží minimálně 24 hodin.

2.3.7 Ochrana před útoky

Datacentrum používá zařízení Radware DefensePro. Zařízení Radware DefensePro jsou známá účinnou ochranou před síťovými útoky v reálném čase. Nežádoucí provoz okamžitě odfiltrují, aniž by uživatelé zaznamenali jakékoliv narušení.

Speciální hardwarová platforma detekuje a umožňuje ochranu před útoky o síle několika desítek Gbps. Chrání tak aplikační infrastrukturu proti výpadkům sítě a pádům aplikací, volumetrickým i flood útokům přicházejícím po síti.

Patentově chráněná behaviorální analýza dat dělá z Radware DefensePro unikátní nástroj. Díky permanentní analýze síťového provozu dokáže v reálném čase detekovat i nové a sofistikované útoky. To vše bez nutnosti zásahu obsluhy a bez narušení legitimního provozu.

2.4 Poskytující služby MasterDC

Firma Master Internet, která vlastní 2 datová centra v Brně a v Praze, nabízí spoustu služeb, které by se daly očekávat od každého většího datacentra. Některé z těchto služeb zmíním i později, v rámci návrhu řešení.

2.4.1 Cloud / VPS

- Cloud hosting
- VPS hosting
- OpenStack cloud
- Živý server
- Private cloud
- Managed servery

2.4.2 Fyzické servery

- Dedikované servery
- Managed servery
- Housing serveru
- Rackhousing
- Pronájem DC

2.4.3 Software

- SaaS
- Mobilní aplikace
- Zakázková softwae

- Audit přípustnosti webu

2.4.4 Datový prostor

- Zálohování
- Pronájem diskového pole
- Pronájem diskového prostoru
- Backup to cloud
- Zálohovací systémy

A mnoho dalších služeb v rámci bezpečnosti nebo např. clusterových řešení a private cloudů.

2.5 O aplikaci

Jedná se o aplikaci **Numbo Call Protect & Blocker**. Tato aplikace umožňuje blokovat telefonáty s nabídkami služeb, průzkumy Agentur nebo jakákoliv další uživatelé nechtěná čísla a to tak, že porovnává volané číslo s nechtěnými telefonními čísly v databázi a zablokuje ho nebo o něm poskytne nějaké základní informace.

2.6 Současný stav aplikace

Aplikace běží na cloudovém serveru umístěném na pobočce MasterDC v Praze. Parametry cloudového serveru jsou:

- 2 jádrový CPU
- 8GB RAM
- 20GB SSD
- OS: CentOS Linux release 7.7.1908 (Core) s kernel 3.10.0-1062.1.1.el7.x86_64

Aplikace je napsaná v programovacím jazyku Java a běží na webovém serveru v Apache Tomcatu 7.0.76, s openjdk Javou 1.8.0_191.

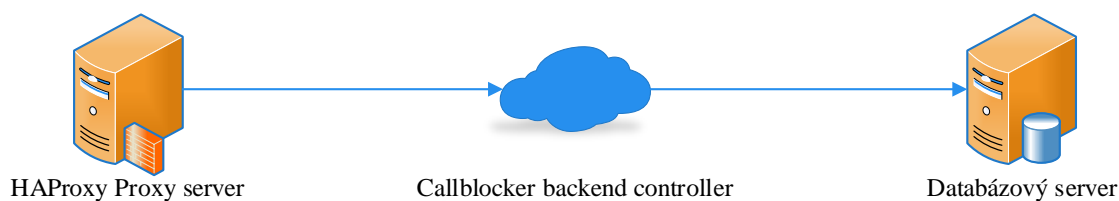
2.7 Aktuální zálohování

Zálohování aplikace je stále ve vývojovém stavu a jeho optimální návrh budu řešit v kapitole návrhu vlastního řešení. Databáze se nachází také na cloudovém serveru, kde zatím běží jen jedna instance databáze.

Parametry cloudového serveru, na kterém je databáze jsou:

- 3 jádrový CPU
- 4GB RAM
- 100GB SSD
- OS: CentOS Linux release 7.7.1908 (Core) s kernel 3.10.0-1062.1.1.el7.x86_64

2.8 Současné zapojení



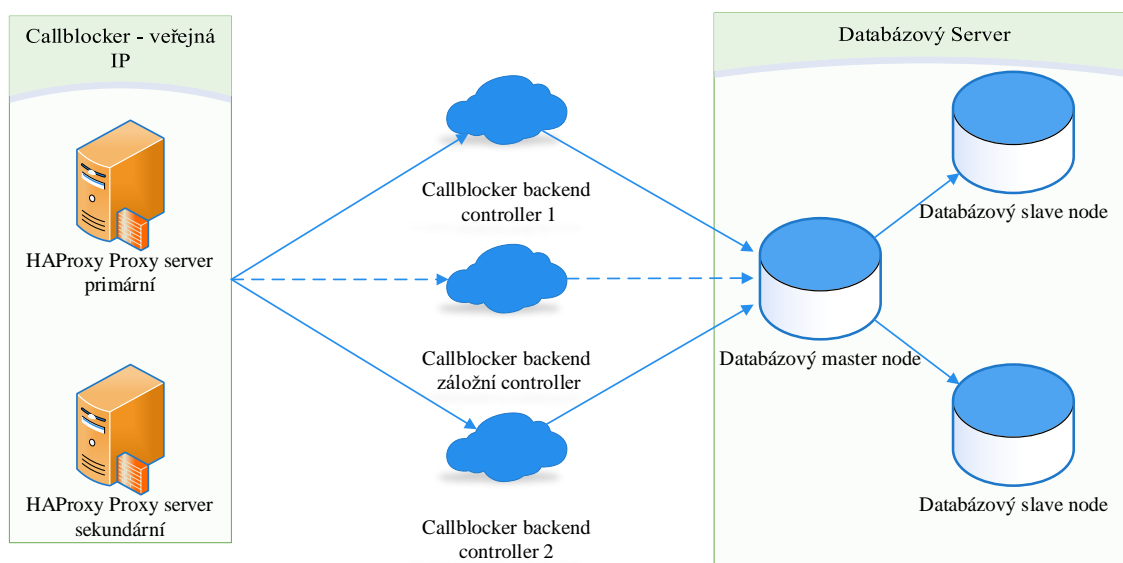
Obrázek 10: Aktuální stav zapojení
(Zdroj: vlastní zpracování)

V současné době je aplikace v takovém stavu, že veřejná IP adresa, na kterou chodí requesty, je na jednom HAProxy Proxy serveru, který je přeposílá na backendový server, jenž vyřizuje takové akce jako registraci, přihlášení, přidání nového zařízení apod. Backendový server poté komunikuje s databází a předá jí request na číslo, které volá zákazníkovi.

3 VLASTNÍ NÁVRHY ŘEŠENÍ

Tato kapitola obsahuje návrh řešení zálohy dat a optimalizace zápisu/čtení dat z databáze pro aplikaci **Numbo Call Protect & Blocker** od společnosti MasterAPP. Vycházel jsem zde z vědomostí mých, nebo nově nabytých při zpracovávání této práce a při konzultaci se zaměstnanci společnosti. Hlavním obsahem této kapitoly bude návrh na zlepšení schopnosti aplikace přistupovat do databáze a vytvoření adekvátní zálohy databázového serveru na základě požadavků a dostupných prostředků společnosti. Poté budou zmíněny přídatné možnosti zálohování k mému návrhu řešení.

3.1 Návrh nového zapojení



Obrázek 11: Vlastní návrh zapojení
(Zdroj: vlastní zpracování)

3.2 Přidání dalšího HAProxy proxy serveru

Dle mého názoru je zde potřeba přidat další HAProxy proxy server, z důvodu stálé dostupnosti veřejné IP adresy. Veřejná VIP IP adresa, přes kterou aplikace přistupuje k backend serverům a následně k databázi, se pomocí routing softwaru Keepalived a protokolu Virtual Router Redundancy Protocol (VRRP) přepíná mezi dvěma HAProxy proxy servery tak, aby byla vždy dostupná i při aktualizaci či výpadku primárního HAProxy proxy serveru.

3.3 Přidání dalších controllerů

Rozhodl jsem se pro přidání dalších 2 backend controlerrů, a to z důvodu redundance a umožnění load balancingu (vyvažování zátěže). S možným počtem růstu uživatelů aplikace, by jeden controller mohl mít problém se zpracováváním requestů na databázi, proto bych navrhoval přidat jeden další aktivní controller, který by měl pomoci zajistit plynulý chod aplikace pro všechny uživatele. Dále bych přidal třetí controller, který by byl neaktivní a sloužil jako redundance v případě, že by se s těmi hlavními dvěma controllery něco stalo, nebo v případě že by oba hlavní controllery nestíhaly.

Přidané controllery budou řešeny pomocí cloud serverů, se stejnými specifikacemi, jako ten, na kterém aplikace již běží, tedy:

- 2 jádrový CPU
- 8GB RAM
- 20GB SSD
- OS: CentOS Linux release 7.7.1908 (Core) s kernel 3.10.0-1062.1.1.el7.x86_64

Pro virtualizaci cloud serverů je použita technologie VMware.

3.4 Přidání více instancí (nodů) databáze

V rámci zálohování a optimalizace databáze jsem navrhl vytvoření 3 databázových nodů, které by fungovaly jako master – slave architektura, a to hned z několika důvodů.

3.4.1 Zálohy

Pro použití replikace jako řešení pro zálohování, je potřeba replikovat data z masteru na slave a poté zálohovat data ze slave nodu. Slave lze pozastavit a vypnout, aniž by to ovlivnilo běžící provoz masteru, takže lze vytvořit efektivní snímek „živých“ dat, která by jinak vyžadovala vypnutí masteru.

3.4.2 Scale-out

Replikace můžeme využít i jako scale-out řešení. To znamená, že lze rozdělit zatížení databázových dotazů na více databázových serverů v rámci určitých rozumných omezení. Protože replikace funguje od distribuce jednoho masteru k jednomu nebo více slaves, použití replikace pro scale-out funguje nejlépe v prostředí, kde máte vysoký počet čtení a nízký počet zápisů / aktualizací, což je přesně případ této aplikace.

3.4.3 Zvýšení výkonu

Se zvyšujícím se počtem slave nodů připojených k masteru se zvyšuje také zátěž, i když minimálně, protože každý slave node musí být neustále připojen k master nodu. Protože každý slave musí obdržet úplnou kopii binárního logu mastera, může se také zvýšit zatížení sítě na masteru a vytvořit takzvaný bottleneck. Při použití většího množství slaves připojených k jednomu masteru a tento master je také zaneprázdněn zpracováním požadavků (například jako součást řešení Scale-out), je možné zlepšit výkon tohoto propojení. Jedním ze způsobů, jak zlepšit výkon procesu replikace, je vytvořit hlubší strukturu replikace, která umožňuje masterovi replikovat se pouze na jeden slave node a zbývající slaves se mohou připojit k tomuto primárnímu slave nodu pro své individuální požadavky na replikaci. Toto řešení by se dalo využít, pokud by bylo potřeba do architektury v budoucnu přidat další slaves.

3.4.4 Převzetí master funkce při selhání

Lze nastavit mastera a slave nody tak, aby slave převzal master funkci v případě potřeby. Stačí napsat skript, který monitoruje master node a kontroluje, zda je aktivní. Poté se

v případě selhání master nodu přikáže nějakému slave nodu, aby převzal funkci mastera, a pokračoval v jeho funkci do té doby, než se původní master obnoví do aktivního stavu.

3.4.5 Zabezpečení

Pro šifrování přenosu binárního logu vyžadovaného během replikace lze použít SSL, ale master i slave toto připojení musí podporovat. Pokud některý node nepodporuje připojení SSL, není replikace prostřednictvím připojení SSL možná. Nastavení replikace pomocí připojení SSL je podobné jako nastavení serveru a klienta pomocí SSL. Musíte získat (nebo vytvořit) vhodný bezpečnostní certifikát, který můžete použít na masteru a podobný certifikát (od stejné certifikační autority) na každém slave.

3.5 Software

Pro zálohování celé databáze jsem po konzultaci se sysadminem zvolil aplikaci BackupPC, jelikož server běží na Linuxové distribuci CentOS a tato aplikace je jedna z nejspolehlivějších v Linuxovém prostředí.

3.5.1 BackupPC

Aplikace BackupPC je ideální na zálohování souborů/složek z jiných strojů na pevný disk. Pro komunikaci se zálohovaným strojem používá protokoly SSH, rsync, nebo SMB. BackupPC samozřejmě také nemá problémy se zálohováním lokálních souborů. Pro samotný přenos pak aplikace využívá protokol rsync, který snižuje nároky na síť i datová úložiště. Zálohy samotné lze iniciovat buď ručně, nebo prostřednictvím zálohovacího plánu. Komprese i odstranění přebytečných záloh je rovněž automatizováno poměrně sofistikovaným způsobem. Výhodou aplikace je, že webové rozhraní, pro příjemnější zaházení a lehkou, přehlednou konfiguraci.



Obrázek 12: BackupPC - hlavní strana
(Zdroj: vlastní)

3.5.2 Instalace

Ve většině distribucí jde balíček stáhnout pomocí package managera dané distribuce. Jeho součástí je rovnou i konfigurační soubor pro Apache a související webové rozhraní. Seznam závislostí pak krom Apache a Perlu doplňuje samba-client, a pár elementárních knihoven pro komprimaci. Ke konci instalace nás aplikace vyzve k zadání hesla, které budeme používat pro přístup přes webové rozhraní. Tím je instalace hotová. Konfiguraci lze pohodlně provádět přes web z adresy <http://localhost/backuppc>, kde se po přihlášení (uživatel backuppc, zvolené heslo) objeví souhrnné info o stavu služby. Nyní je vhodné nastavit obecné volby, které se pak použijí jako výchozí pro každý zálohovaný stroj.

3.5.3 Obecná konfigurace

Základní nastavení lze najít vlevo pod položkou Edit Config. V hlavní části je pak pro přehlednost rozdělena na jednotlivé záložky. Jsou jimi:

- Hosts – Seznam zálohovaných strojů a uživatelů webového rozhraní s povoleným přístupem.
- Xfer – Nastavení přenosu mezi zálohovaným a zálohovacím strojem, které soubory zálohovat a které ne.

- Email – Nastavení kontaktních adres, četnosti informování a předloh jednotlivých zpráv.
- CGI – Nastavení webového rozhraní, aktivace přídatných modulů.
- Server – Cesty k programům, omezení zpracování ve více vláknech.
- Backup Settings – Kde a jak hledat zálohované stroje.
- Schedule – Plánování vytváření a mazání záloh.

BackupPC

Hosts

Select a host.....

Go

Server

[Status](#)

[Host Summary](#)

[Edit Config](#)

[Edit Hosts](#)

[Admin Options](#)

[LOG file](#)

[CGI LOGs](#)

[Email summary](#)

[Current backups](#)

[Documentation](#)

[Wiki](#)

[SourceForge](#)

Main Configuration Editor

Save

[Hosts](#) [Xfer](#) [Email](#) [CGI](#) [Server](#) [Backup Settings](#) [Schedule](#)

Xfer Settings

XferMethod: rsync

XferLogLevel: 1

ClientCharset:

ClientCharsetLegacy: iso-8859-1

Rsync Settings

RsyncShareName: Insert / Add

RsyncCaumCacheVerifyProb: 0.01

Include/Exclude

Path	Insert	Delete
/boot	Insert	Delete
/etc	Insert	Delete
/opt	Insert	Delete
/var/lib/portage/world	Insert	Delete
/var/spool/mail	Insert	Delete
/var/lib/mailman	Insert	Delete
/var/www	Insert	Delete
/home/printex	Insert	Delete
Add		

New Key: Add

Rsync Paths/Commands/Args

RsyncClientPath: /usr/bin/rsync

RsyncClientCmd: sshPath -q -x -l root \$host \$rsyncPath \$argList+

RsyncClientRestoreCmd: sshPath -q -x -l root \$host \$rsyncPath \$argList+

Arg	Insert	Delete
--numeric-ids	Insert	Delete
--perms	Insert	Delete
--owner	Insert	Delete
--group	Insert	Delete
-D	Insert	Delete
--links	Insert	Delete

Obrázek 13: BackupPC - nastavení parametrů zálohování
(Zdroj: vlastní)

3.5.4 Přidání stroje a konkrétnější nastavení

Do správy strojů se lze dostat z levého menu odkazem Edit Hosts. Zde lze připsat název stroje a seznam. Ve výchozím stavu bude tento stroj vyhledán prostřednictvím DNS, kterému bude jako dotaz zasláno jméno stroje. Toto lze však změnit zadáním adresy (IP nebo DNS) v konfiguraci hosta (Backup Settings | ClientNameAlias). Předvyplněné volby dávají najevo, že všechny konfigurační volby byly přejaty z nastavení serveru. Každou je ale možné změnit dle chuti.

3.6 Zálohování Databázového serveru

Aplikace BackupPC bude nastavena tak, aby držela 1 plnou a 4 inkrementální zálohy serveru, což zajistí dostatečnou úroveň zálohování. V záložce Schedule webového rozhraní aplikace se pak nastaví spuštění na 22:00, aby se uložily veškeré změny z průběhu dne.

Následně se bude používat zálohovací skript (viz. příklad níže), který bude dělat dump databáze a ukládat ho do určené složky. Skript se bude používat přes Cron, což je linuxová utilita, která umí spustit nějaký proces, v tomto případě skript, v určitý čas. Cron tento skript bude používat každý den v 1 hodinu ráno, kdy se předpokládá nejmenší aktivita databáze a systém bude v těchto hodinách také nejméně zatížen. Skript se bude provádět na slave instanci databáze, aby se nenarušil chod master databáze.

Databázový server je poté opět zálohován pomocí aplikace BackupPC a ze zálohy budou odstraněny nepotřebné složky typu /dev, /proc, /tmp atd. pro urychlení zálohy a šetření místa.

3.7 Zálohovací skript

Zde jsem ve spolupráci se sysadminem společnosti navrhnul skript, o kterém jsem mluvil výše. Skript bude provádět sql dump do určené složky. Sql dump se bude provádět po tabulkách, jelikož je to přehlednější a lépe se potom řeší případné opravy, kdyby zrovna nastala nějaká chyba v databázi. V případě sql dumpu celé databáze by se totiž v případě

nějaké chyby musel udělat celý znovu, a to je při velikosti této databáze jak časově, tak systémově velice náročné.

```
#!/bin/bash
#
# MySQL Backup Script by tables
#
#####
# Configuration
#
directory=$(date +"%Y-%m-%d")
path="/backup/mysql"
old=$(date --date="14 days ago" +"%Y-%m-%d")

#####
# Script
#
echo "mySQL Backup Script"

# test connection to mysql
TESTCON=`echo "SELECT 1=1" | mysql 2>&1`
if [[ $TESTCON == *"Access denied"* ]]; then
echo "Connection to mysql failed"
exit 1
fi

# create destination directory
mkdir -p $path
mkdir -p $path/$directory

echo "MySQL backup by tables:"

# cp ibdata* ib_log*
#echo "=> Backup ib* files"
#find /var/lib/mysql -iname "ibdata*" -exec cp '{}' $path/$directory/ ';'
#find /var/lib/mysql -iname "ib_logfile*" -exec cp '{}' $path/$directory/ ';'

```

```

# dump entire DB
#echo "=> Dumping entire databases"
#mysqldump --add-drop-table --allow-keywords --all-databases > $path/$directory/all.sql

# dump by tables
echo ""
echo "=> Dumping individual tables"

for a in `echo "show databases" | mysql | grep -v Database | grep -v information_schema`;
do
    mkdir -p $path/$directory/$a
    echo "-> Dumping database: $a"

    for i in `echo "use $a;show tables" | mysql $a | grep -v Tables_in_`;
    do
        echo " * Dumping table: $i"
        mysqldump --add-drop-table --allow-keywords -R -q -a -c $a $i >
$path/$directory/$a/$i.sql
    done
done

# make archive file
echo "=> Archiving Files"

# make zip
origin_path=`pwd`
cd $path; rm -rf "$directory.zip"; zip -r $directory $directory; cd $origin_path

# make tar.gz
#tar -C$path -c -f$path/$directory.tar $directory
#gzip -f $path/$directory.tar

# delete directory and old files
echo "=> Delete directory"
rm -rf $path/$directory
echo "=> Delete archve from $old"
rm -rf $path/$old.*

chmod og+rx /backup

```

```
chmod og+rx /backup/mysql
```

```
# done
```

```
echo "=> Done"
```

tento skript dumpuje pomocí mysqldump databázi po tabulkách do složky /backup/mysql, která je určena pomocí proměnné \$path na začátku. Následně vše komprimuje přes zip a promazává soubory starší 14 dnů.

3.8 Cloudové úložiště

Na cloudová úložiště se dnes zálohuje zcela běžně. V první řadě je tato záloha bezpečná vůči fyzickému poškození, způsobenému například vytopením, požárem nebo i lidským faktorem. Také je to velmi jednoduché, rychlé a relativně finančně nenáročné řešení.

Pro možnost cloudové zálohy lze využít službu Managed backup přímo nabízenou datovými centry MasterDC ve vlastnictví společnosti, spolu s využitím cloudového úložiště, které nabízí taktéž tato datová centra. Cloud hosting společnosti je postaven na nejmodernějších Dell EMC technologiích. Jelikož databázový server je umístěn na pobočce v Brně, cloudové úložiště by bylo umístěno na pobočce v Praze, tímto se zajistí takzvaná offsite záloha.

V rámci služby Managed backup se provádí jedenkrát denně rozdílová záloha a poté jednou za 14 dní záloha plná. Pro linuxové servery se zde opět používá aplikace BackupPC, kterou jsem představil dříve. Je zde také možnost přístupu přes webové rozhraní, odkud jde provádět samoobslužná obnova dat bez nutnosti kontaktovat technickou podporu.

3.9 Záloha pomocí snapshotů

Společnost ve svých službách také nabízí možnost zálohy pomocí databázových snímků (snapshotů). Největší předností tohoto zálohování je jeho rychlost a spolehlivost. Kompletní obnova dat se zkrátí z hodin řádově na minuty a díky snapshotům navíc nedochází ke změně dat mezi počátkem a koncem procesu zálohování.

Snímek databáze nabízí virtuální konzistentní kopii databáze pouze pro čtení. Když vytváříme snímek databáze v živé operační databázi, zabere to statický pohled na databázový bod v čase a zpětné vrácení všech nesouvisejících transakcí v databázi snímků, takže nebudeme mít žádná nekonzistentní data, která ještě musí být potvrzena.

V tomto řešení lze databázový snímek použít například před nějakým větším updatem nebo plánovaných změnách v databázi. Takto lze databázi při nějaké chybě, ať už lidské nebo systémové, snadno vrátit do původního stavu, tedy stavu, kdy byl snímek pořízen a předejít tak složitému opravování databáze. Jelikož databázový server je vytvořen virtuálně, lze tu také udělat snímek celého systému k určitému času. To má samozřejmě jako velkou výhodu také rychlost, oproti běžné záloze celého systému, která je většinou zdoluhavá. Tento snímek systému by se poté mohl zálohovat na cloud v Praze o kterém jsem mluvil výše, a tak sloužit jako plnohodnotná záloha.

3.10 Zhodnocení navrhovaného řešení

Navrhované řešení bylo přizpůsobeno požadavkům zaměstnavatele a dostupným prostředkům firmy. Aplikace tohoto řešení zajistí optimální chod aplikace do budoucna ve spojení s odhadovaným růstem uživatelů. Dále toto řešení poskytuje adekvátní zálohu v souvislosti s velikostí databáze pro aplikaci a bezpečností datacentra s využitím produktů a služeb poskytovaných samotnými datacentry firmy MasterDC.

Hlavní záloha prostřednictvím aplikace BackupPC bude plně automatizována a bude probíhat každý den v nočních hodinách. Záloha na cloudový server nebo pomocí snapshotů lze také plně automatizovat a byla by provedena zkušenými zaměstnanci datacentra.

ZÁVĚR

Cílem této bakalářské práce bylo vytvoření návrhu řešení zálohování dat pro zefektivnění práce s uloženými daty a zajištění jejich vyšší bezpečnosti. K dosažení tohoto cíle jsem měl vypracovat teoretická východiska, která by sloužila jako znalostní podklad pro další části práce, kterými jsou analýza současného stavu a vlastní návrhy řešení. Teoretická východiska jsem zdárně vypracoval, a tak jsem je mohl využít při tvorbě dalších částí práce.

Dále jsem měl provést analýzu současného stavu, kde jsem měl představit společnost, pro kterou návrh řešení dělám, objasnit vybavení firmy a současný stav zálohování aplikace. I tato část byla zdárně vypracována.

Poslední částí práce, která měla být pro dosažení cíle vytvořena byla část vlastní návrh řešení. Zde jsem využil teoretických východisek práce, vlastních znalostí, znalostí nově nabytých z konzultací se zaměstnanci firmy nebo vlastním průzkumem k danému tématu. Ve finále se mi podařilo ve spolupráci se sysadminem firmy vytvořit adekvátní a proveditelný návrh řešení, který odpovídá požadavkům a zajišťuje tak optimální a bezpečnou zálohu dat. Také je tímto řešením zajištěn bezproblémový přístup k databázi z pohledu aplikace, i při větší zátěži ve formě databázových dotazů.

SEZNAM POUŽITÝCH ZDROJŮ

- (1) Jak na Internet - Zálohování. Jak na Internet - Jak na Internet [online]. Copyright © 2020 CZ.NIC, z. s. p. o. [cit. 24.04.2020]. Dostupné z: <https://www.jaknainternet.cz/page/1180/zalohovani/>
- (2) DOSEDĚL, Tomáš. Počítačová bezpečnost a ochrana dat. Brno: Computer Press, 2004. ISBN 80-251-0106-1.
- (3) 6 způsobů, jak zálohovat důležitá data. Použít cloud, NAS server, flash disk, hard disk nebo zálohovací software?. F22.cz - Fotografie nás baví - [online]. Copyright © 2020 [cit. 24.04.2020]. Dostupné z: <https://f22.cz/6-zpusobu-zalohovat-svoje-dulezita-data/>
- (4) Co je NAS server - NAS servery a chytrá domácí úložiště dat. NAS servery - NAS servery a chytrá domácí úložiště dat [online]. Copyright © 2020. [cit. 21.04.2020]. Dostupné z: <http://www.nasservery.cz/je-nas-server/>
- (5) Storage Area Network – 1 (úvod). AbcLinuxu.cz - Linux na stříbrném podnose [online]. Copyright © 1999 [cit. 12.04.2020]. Dostupné z: <https://www.abclinuxu.cz/clanky/storage-area-network-1-uvod>
- (6) What Is a Storage Area Network (SAN)? | SNIA. SNIA | Advancing Storage and Information Technology [online]. Copyright © 2020 SNIA [cit. 21.04.2020]. Dostupné z: https://www.snia.org/education/storage_networking_primer/san/what_san
- (7) CONRAD, Eric, Seth MISENAR, Joshua FELDMAN a Kevin RIGGINS. Eleventh hour CISSP: study guide. Second edition. Amsterdam: Elsevier, 2014. ISBN 01-2417142-7.
- (8) Inkrementální – přípustková záloha. acronis.cz [online]. [cit. 22.04. 2020]. Dostupné z: <https://www.acronis.cz/kb/inkrementalni-zaloha/>

- (9) Diferenciální - rozdílová záloha. acronis.cz [online]. [cit. 22.04. 2020]. Dostupné z: <https://www.acronis.cz/kb/diferencialni-zaloha/>
- (10) [online]. Dostupné z: <https://it-slovník.cz/pojem/raid>
- (11) RAID Levels Explained | PCMag. The Latest Technology Product Reviews, News, Tips, and Deals | PCMag [online]. Copyright © 1996 [cit. 24.04.2020]. Dostupné z: <https://www.pcmag.com/news/raid-levels-explained>
- (12) Co to je RAID a k čemu slouží? | GIGA PC. GIGA PC - Specialisté na repasované počítače [online]. Dostupné z: <https://www.giga-pc.cz/technicke-okenko/raid/>
- (13) What is Database Clustering - Introduction and brief explanation - NdimensionZ. 301 Moved Permanently [online]. Copyright © 2020. All Rights Reserved. [cit. 28.04.2020]. Dostupné z: <https://ndimensionz.com/kb/what-is-database-clustering-introduction-and-brief-explanation/>
- (14) MySQL Cluster Performance Validation via Load Testing | Continuent Ltd. Continuent Ltd | MySQL Availability | MySQL Cloud | MySQL Cluster | Geo-Scale | Zero-Downtime | MySQL Replication [online]. Dostupné z: <https://www.continuent.com/blog/mysql-cluster-performance-validation-via-load-testing/>
- (15) Datacentrum MasterDC v Praze a v Brně | Master.cz. Server Hosting, Housing, Virtuální servery VPS - Master Internet [online]. Copyright © 2020 Master Internet, s.r.o. [cit. 01.05.2020]. Dostupné z: <https://www.master.cz/datacentrum/>

SEZNAM POUŽITÝCH OBRÁZKŮ

Obrázek 1: Důvody ztráty dat.....	13
Obrázek 2: NAS	16
Obrázek 3: Druhy záloh.....	18
Obrázek 4: Raid 0.....	20
Obrázek 5: Raid 1.....	21
Obrázek 6: Raid 5.....	22
Obrázek 7: Raid 6.....	23
Obrázek 8: Databázový cluster	24
Obrázek 9: Propojení datacenter	28
Obrázek 10: Aktuální stav zapojení	32
Obrázek 11: Vlastní návrh zapojení.....	33
Obrázek 12: BackupPC - hlavní strana	37
Obrázek 13: BackupPC - nastavení parametrů zálohování	38

SEZNAM POUŽITÝCH GRAFŮ

SEZNAM POUŽITÝCH TABULEK